

光音网络的 存储容器化方案探索

王鹏 (Tad)

Container如何落地?

传统项目容器化的困境

硬件多样性

- 3U: 1 * SSD, 4 * HDD
- 6U: 1 * SSD, 1 * HDD
- 4U: 36 * HDD, 72 * HDD
- ...

多种操作系统



数据相关系统



elasticsearch.



GlusterFS



redis



MySQL

kafka



mongoDB



MariaDB



hadoop



Apache ZooKeeper™



etcd



ceph



NSQ

数据支撑的一系列问题

- 传统项目容器化的困境
- 抗并发能力
- 动态扩缩容
- ...

Containerize

Why Container?

分类

- 分布式文件系统
- 数据库系统
- 无状态的本地系统

探索的内容

- 部署
- 升级
- 扩容及负载均衡
- 灾备及备份
- 基准测试

部署

常规部署的问题

- 下载慢
- 安装复杂
- 各种版本
- 各种数据库
- 操作系统的不同
- 宿主机环境的依赖
- 各种不同的调优参数

如何简化

- Ansible
- 服务器统一资源池
- 统一部署模板 (start stop)
- 监控配套

如何简化

Monitor

Ansible

Template

Docker Image

Server Resource Pool

Template

```
Description=mongo-crawler-rs
After=docker.service
Requires=docker.service
[Service]
TimeoutStartSec=0
Restart=always
ExecStartPre=-/usr/bin/docker rm mongo-crawler-rs
ExecStart=/usr/bin/docker run \
  --name mongo-crawler-rs \
  --net:host \
  -v /hdd1/crawler-rs:/data/db \
  PRIVATE-REGISTRY/mongo:3.0 \
  --storageEngine wiredTiger \
  --replSet crawler-rs
  -- keyFile=/data/db/keyFile

ExecStop=/usr/bin/docker kill mongo-crawler-rs
[Install]
WantedBy=multi-user.target
```


如何创建



Create Mongo

Name

crawler-mongo

Replica Set Amount

3

Shard Amount

5

Operations / Second

10000

Storage Size (GB)

100

Clients Amount

20

Service Domain

adx.mongo.cloud.localhost

CREATE

升级

优势

- 试升级操作
- 回滚版本操作
- 环境依赖问题
- 不同版本的数据库共存
- 复杂的升级安装过程

Example

```
Description=mongo-crawler-rs
After=docker.service
Requires=docker.service
[Service]
TimeoutStartSec=0
Restart=always
ExecStartPre=-/usr/bin/docker rm mongo-crawler-rs
ExecStart=/usr/bin/docker run \
  --name mongo-crawler-rs \
  --net=host \
  -v /hdd1/bds-uba-rs:/data/db \
  PRIVATE-REGISTRY/library/mongo:3.2 \
  --storageEngine wiredTiger \
  --replSet crawler-rs
  -- keyFile=/data/db/keyFile

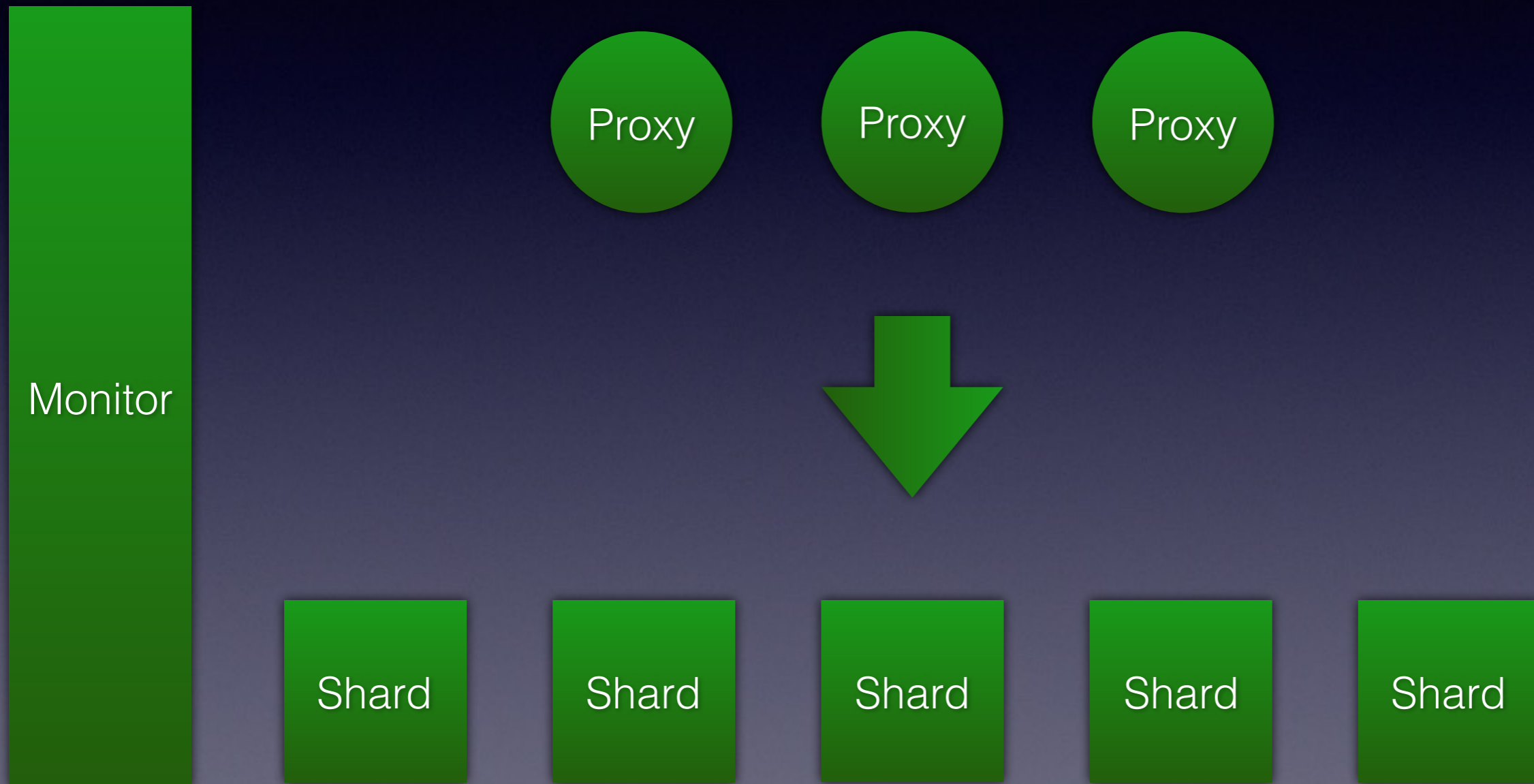
ExecStop=/usr/bin/docker kill mongo-bds-uba-rs
[Install]
WantedBy=multi-user.target
```

然并升

- 小版本
 - 支持平滑升级
- 大的引擎变化
 - 副本集的同步
 - 数据的重导入
 - 升级脚本

水平扩展

扩展及自动负载示意



Shard Example

```
[Unit]
Description=mongo-shard
After=docker.service
Requires=docker.service
[Service]
TimeoutStartSec=0
Restart=always
ExecStartPre=-/usr/bin/docker rm mongo-shard
ExecStart=/usr/bin/docker run \
  --name mongo-shard \
  --net=host \
  -v /data/shard:/data/db \
  PRIVATE-REGISTRY/library/mongo:3 \
  --storageEngine wiredTiger \
  --shardsvr
ExecStop=-/usr/bin/docker kill mongo-shard
[Install]
WantedBy=multi-user.target
```


数据支持列表

 kafka



elasticsearch.



GlusterFS



mongoDB



redis



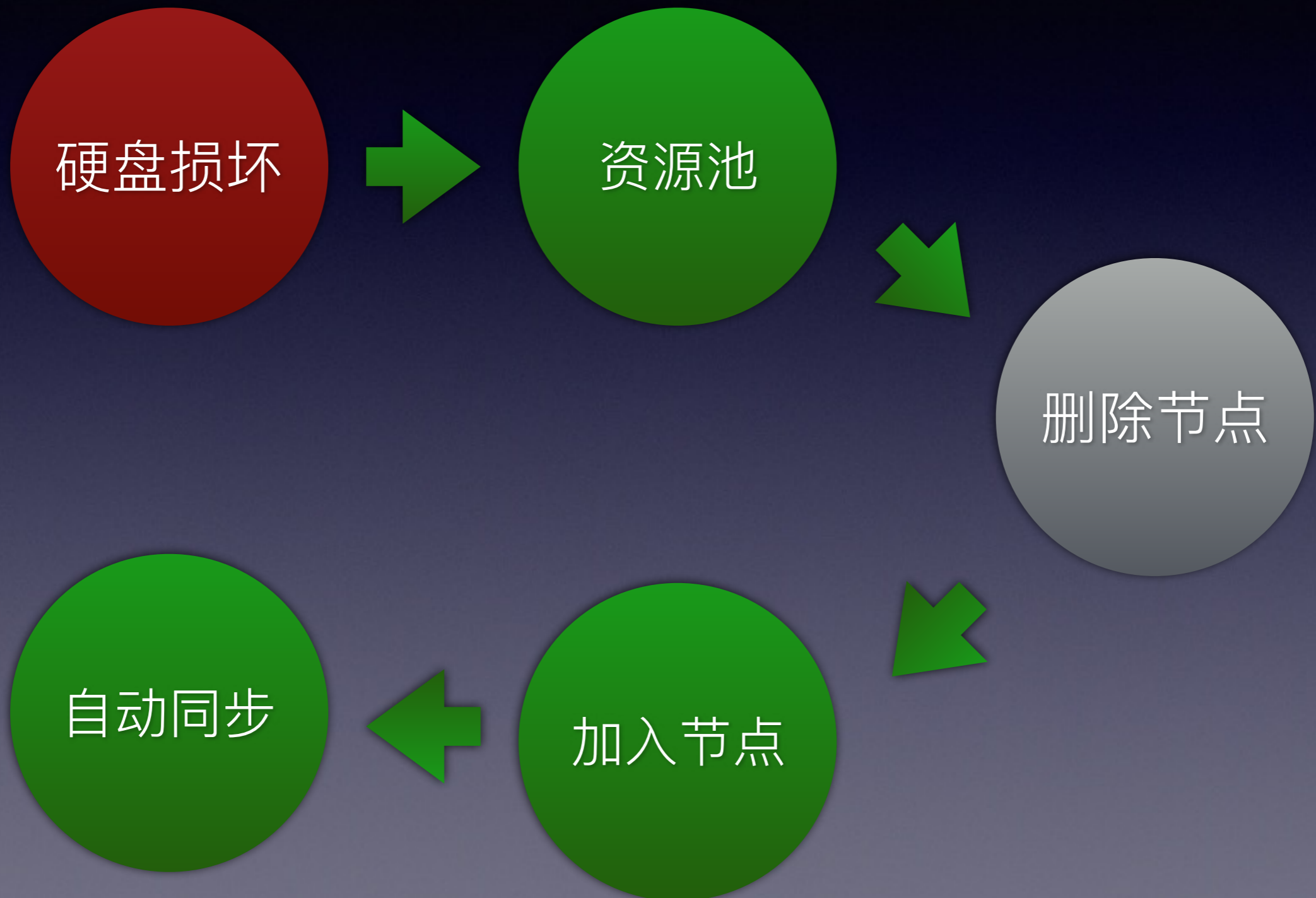
ceph



NSQ

容灾及备份

容灾操作示例



数据支持列表



elasticsearch.



GlusterFS



redis



MySQL®

kafka



mongoDB



MariaDB



Apache ZooKeeper™



etcd



ceph



NSQ

备份

- 副本集
- 冗余
- 延时同步
- 增量备份

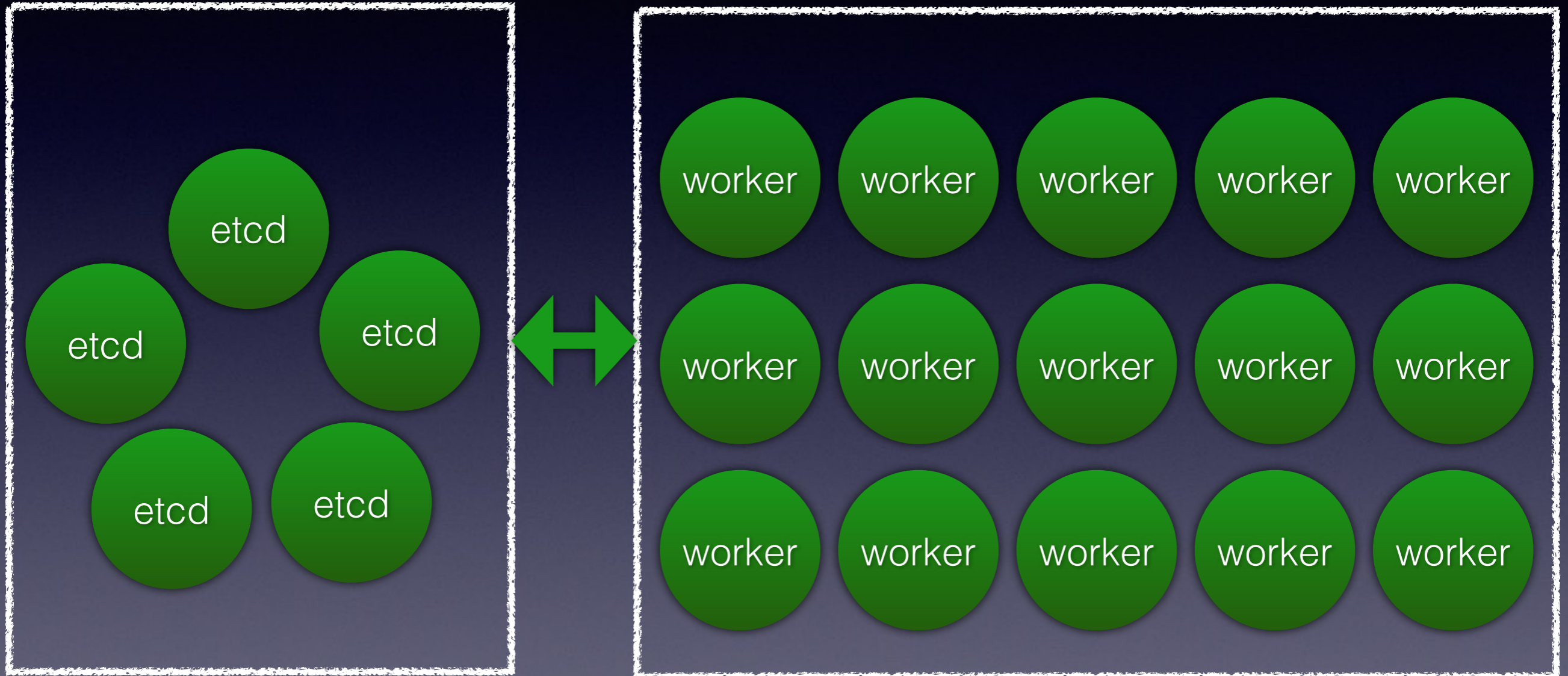
备份监控—配套执行

基准测试

优势

- 充分利用闲时的服务资源
- 模拟真实业务更大的并发
- 每次版本测试的用例一样
- 构建一次，可以终生使用

调度示例



容器化的优势

- 部署升级的统一化操作 (OS, Operations)
- 充分利用存储资源 (包括硬盘, CPU, 内存)
- 调优的一致性
- 动态地扩缩容

归一化

踩过的坑

踩过的坑

- 全过程自动化
- Ceph VS GlusterFS
- 分小区，做隔离，降低风险
- [facebook/dvara](https://www.facebook.com/dvara)

Thank You

We Are Hiring!

<http://www.goyoo.com>

